Analysis Of Engaged Online Collaborative Discourse: A Methodological Approach

Sofia Malheiro; Lina Morgado; António Quintas-Mendes
Universidade Aberta, Portugal
sofiamalheiro@gmail.com
lmorgado@univ-ab.pt
quintas@univ-ab.pt

I. Introduction

The objective of this chapter is to reflect upon a new educational field which involves computer-mediated collaborative learning, by systematizing the main methodological constraints we face when we seek to analyze the interactions taking place in that collaboration.

In fact, to deepen the analysis of student participation in discussion forums integrated in online distance training courses is of the utmost importance in order to produce strong evidence about the learning taking place and the construction of knowledge that might occur.

As Fahy reminds us [2], researchers should be able to describe online interactions in a less impressionistic way and measure them more efficiently. We can thus consider some studies in which transcriptions of the computer-mediated conferences were examined in a more precise manner.

Henri [3] centered his analysis on the learning processes perceived in the messages. His model grants particular emphasis to content analysis.

Gunawardena [11] took Henri's model further and shifted the emphasis to the nature of the dialogue taking place, highlighting the importance of the analysis of interaction itself.

Lally [4], taking into account Henri's content analysis and Gun-

tive learning environments "provide means to evaluate and confirm the nature of the existing complex relations between teaching and learning which are not accessible in conventional classrooms".

In our readings so far there seems to be a consensual view that online conferences have, at their core, a learning environment that supports the notion of "knowledge construction" in a social framework.

The messages in computer-mediated conferencing are widely acknowledged as a new, hybrid form of communication. They contain some of the strengths and weaknesses of speech and of writing, but it is not quite clear how good they are to make the most of those strengths and minimize the weaknesses in online learning environments.

Many studies have resorted to techniques which, developed from measures of syntactic or linguistic analysis, among others, indicate the construction of collaborative meaning or joint interaction.

At present, there is no clear perspective on the best way to analyze these characteristics of virtual communication.

In her research, Xin [11] tried to create an interpretative framework which allowed for measuring the quality of engaged collaborative discourse (ECD) in computer conferencing. To that end, she used three sub-scales that measure, respectively, the Communication Processes, the Intellectual Engagement and the Use of Moderating Functions in ECD.

These three sub-scales will then be combined into a single scale that aims at measuring the Overall Quality of Engaged Collaborative Discourse.

This is the conceptual frame in which we conducted an empirical study, having as object an Online Trainers Training Course offered by Universidade Aberta.

This work systematizes the main methodological constraints we faced when we used Xin's scales to analyze the level of engaged collaborative discourse in a discussion forum, part of the module "Models of Online Training: Collaborative Learning", with the objective of determining the nature of the interactions that took place.

II. Conceptual Framework

1. Distance Education and the Internet

One of the basic characteristics of Distance Education is the establishing of an interactive communication between teachers and students. Because they don't share the same physical space, they need means that enable the communication between them.

If we wish to analyze the use of technology in distance training we can identify a clear cleavage: distance training *before* and *after* the emergence of the Internet. The predominant technologies in the former generation of

distance education allowed one-to-one (telephone) and one-to-many (television) communication. The Internet brought a new way to develop the teaching and learning process by supporting a new form of distance interaction - many-to-many – that can be achieved through several communication resources (e-mail, chat, discussion forums, etc.) which have to potential to enable cooperative and collaborative learning.

As new competencies to deal with the world of information and knowledge emerge, new models of teaching and learning arise, bringing forth some reformulations in the conventional training processes and generating a need for change in the established educational paradigms.

Running parallel to these technological advancements, the pedagogical consequences of their use have brought us to a transition phase in the courses offered, from a more self-instructional nature to a more collaborative/interactionist one.

2. Collaborative Learning and Cooperative Learning

The concept of the teaching and learning process appears frequently in the literature on pedagogy or didactics. For many centuries there was a strong belief that knowledge was something fluid that could be transmitted by the teacher to the learner. In other words, there was the belief that it is possible for a teacher to teach a student (i.e. transmit knowledge, information or explanations which are useful or indispensable to education and instruction). This still is the epistemological basis for education in the majority of the contemporary institutions and at all levels of education.

The scientific studies conducted in the last century by Jean Piaget on how the human being acquires knowledge point in a different direction. Piaget considers that cognitive development occurs in the interaction of a subject and an object. The cognitive subject functions on the basis of signification schemes which undergo permanent *adaptation* through *assimilation* processes, continuous and simultaneous (the subject's schemes are modified to incorporate the object's elements), and *accommodation* processes (the object's elements are modified by the subject's action). Thus, cognitive growth requires that a disequilibrium occurs in this process, which will cause the appearance of new schemes based on the existing ones, triggering a spiral connected to many others through individual signification webs. In this context, learning emerges as an individual process that takes place internally to the subject.

Vygotsky [12], on the other hand, postulated that the higher psychological processes appear firstly in social relations in the form of interpersonal processes, then becoming intrapersonal or individual.

The justification for the social origin in the higher psychological processes lies, according to Vygotsky, in the mediation performed by tools. These tools can be physical (tools that control the environment) or psychological (signs, language in particular), and the relation of the human being with them is active and transformative.

The move from the interpersonal (social) to the intrapersonal (individual) process happens through interiorization, which is "the internal reconstruction of an external operation" [12:74]. Both oral and written language are made of symbols used in human communication which, when interiorized, create new forms of thinking. The interiorization of socially rooted and historically developed activities constitutes the basic feature of human psychology [12:76].

Learning is thus a social process that takes place through the possibilities created by the subject's mediations in a given social and historical context, since, as Ratner [12:160] points out, "the individual does not face things with a solitary conscience. He is a member of a social community and depends on other people for material, behavioral and psychological help". We therefore reassert Vygotsky's conviction that social interaction is the origin of and the driving force behind learning and intellectual development [12].

Pierre Lévy [5] also incorporates the sociocultural approach to learning when he proposes the virtual learning communities, the collective construction of shared knowledge and the networked collaborative/cooperative learning. To learn with others, reformulating knowledge through the critical perspective of others, is important to strengthen reasoning and communication skills. The fundamental assumption of the collaborative/cooperative models is that the acquisition and development of knowledge, skills or attitudes is not an inherently individual process but rather results from group interaction.

In the context of his Cognitive Ecology theory, Pierre Lévy [4] states:

> "intelligence or cognition are the result of complex networks where a great number of human, biological and technical actors interact. It is not 'I' who am intelligent but 'I' with the human group of which I am a member, with my language, with a whole heritage of intellectual methods and technologies" (p.135)

Piaget, who was criticized for not having favored the social element in cognitive development, also considered the importance of interactions. For him [9] "it is precisely the constant interchange of thoughts with others that allows us to decenter and assure the possibility of coordinating internally as relations that emanate from different points of view". In his book *Estudos Sociológicos* (Sociological Studies) [9], Piaget tries to find a sociological explanation, as opposed to a psychological explanation of cogni-

tive growth. The basic concept of the subject's action upon the object remains, with the difference that the "I" has now become "We" and that the actions give way to interactions or "conducts that modify one another, forms of cooperation, that is, operations performed in common or in reciprocal correspondence" (p. 22).

Continuing Piaget's work, Perret-Clermont [8] concentrated on the influence of interactions on cognitive development, using the results of the studies conducted by Doise & Mugny [cit 8] which showed that, under certain circumstances, pair work was better than individual work. It follows from these considerations that learning takes place inside each individual, but can be prompted by social exchanges, i.e. social interactions and culture play, in this perspective, a prominent role in the individual's cognitive development. This is the basic principle of Cooperative/Collaborative Learning, which in this case follows the socioconstructivist approach.

There is, however, some controversy about the use of the terms "cooperative learning" and "collaborative learning". The debate of "collaboration" versus "cooperation" is a complex one. It is possible to state, on a first approach, that both paradigms are based on a constructivist epistemology, with a special relevance to the learner's active role.

In the distinction put forward by Panitz [7], the term "collaboration" corresponds to a "a philosophy of interaction and personal lifestyle where individuals are responsible for their actions, including learning and respect the abilities and contributions of their peers".

The term "cooperation" is viewed, by the same author, as a "structure of interaction designed to facilitate the accomplishment of a specific end product or goal through people working together in groups". Cooperative learning is thus defined by a set of processes that help individuals interact to achieve a common goal.

Dillenbourg [1] also views collaboration as different from cooperation. This author distinguishes the concepts along three main characteristics:

- Degree of symmetry in the interaction
- Shared goals
- Division of labor

For him, situations can be characterized as more or less collaborative or more or less cooperative according to these three criteria.

The situations of symmetry in the interaction (action, knowledge and status), analyzed from the objective and/or subjective point of view, might lead to classify the task as more collaborative or more cooperative. By the same token, the existence or absence of *shared goals* within the group is a

strong indicator that allows us to classify the task as being collaborative or cooperative.

For Dillenbourg [1], the way the different elements in the group perform the *division of labor* also leads us to this distinction. Thus, in cooperation the partners divide labor vertically, in independent subtasks, solve them individually and them put the final work together; in collaboration the partners work together, although there might be some horizontal subdivisions.

It seems therefore relevant to distinguish these two concepts and suggest that what we might term as collaborative learning results from the assumption of principles of solidarity and empathy towards others, without any other explicit obligations; whereas cooperative learning rests on clear principles which regulate the techniques to be used in the group, goals/objectives shared by all members and a fixed, explicit division of the labor of each element in the group.

In this perspective, the concept of cooperation seems to be more complex than those of interaction and collaboration, since it not only presupposes these but also requires non-hierarchical, mutual respect relations among the individuals involved, a posture of tolerance with and acceptance of the differences in a process of constant negotiation, along with goals, activities and actions which are conjoint, coordinated and shared.

To move from the theoretical elaboration on collaboration and cooperation to the empirical analysis of interactions, seeking to understand the ways in which dialogue and cooperation take place online is our strong motivation in this phase of searching for valid tools and methodological procedures to analyze interaction in computer mediated communication.

III. Analysis of Engaged Collaborative Discourse

For Xin [13], to create an interpretative framework that permits the measuring of individual learning and engaged collaborative discourse in online discussions implies the establishing of a process that is synergetic and, simultaneously, allies design and development at the theoretical and instrumental level, allowing for the articulation of some of the variables. Thus, Xin proposes three subscales that after being combined will produce a single scale that measures the overall quality of engaged collaborative discourse (ECD).

1st Subscale – Measurement of the Communication Processes

"Sustained group communication is reflected in its members' contributions to the discourse and their interaction with others" [13:155].

In this perspective, the concept of *participation* is related to contribution and interaction. To measure the "quantity" of communication three rating questions are used, which aim at describing the level of individual participation on a given discussion topic and in a given period of time. They are:

- **Login activities** – Number of times an individual makes a request to download waiting messages from the server in a given time period; number of messages read by an individual at a given point in time.
- **Quantity of contribution** – Number of messages posted by an individual in a given discussion topic and the total number of words in those messages.
- **Quantity of interaction –** Number of references made by an individual to previous contributions in a given time period in a given discussion topic.

The results from these three rating questions allow a classification of the individual according to six levels of communication: Peripheral; Minimal Participation; Basic; Regular; Frequent; Active.

This subscale thus tries to answer the questions of "How much" and "How often" an individual contributes and interacts. However, it does not answer the question regarding the "Quality of participation" of the individuals.

2nd Subscale – Measurement of Intellectual Engagement

To establish a rubric for measuring the quality of intellectual engagement Xin [13] uses three rating questions that try to describe the level of Intellectual Engagement of an individual in a given discussion topic during a given time period. They are:

- **Coherent with the teacher's agenda for the discussion** – The most basic requirement for participation in ECD is to be able to follow the teacher's agenda and post relevant and understandable contributions.
- **New ideas and points of view** – Contributing with new ideas and new points of view are important indicators of intellectual involvement and understanding, and can happen at any time in the discussion.
- **Relation with other contributions** – Elaborating on and relating to one or several previous contributions happens in the negotiation, construction and integration zones.

The answers to these three rating questions allow the classification of the individual according to five levels of intellectual engagement: Poor, Minimal, Fair, Competent, Excellent.

This subscale is not independent from the Communication Subscale. Both point to levels of intellectual contribution and interaction. The first is more focused on the quantitative aspects, the second on the qualitative ones.

3rd Subscale - Measuring the Use of Moderating Functions

This third subscale measures the use of Moderation Functions by the individual – student or teacher/tutor – in a discussion topic, in a given period of time.

Moderation functions describe the tasks to be performed so as to facilitate and support the communication processes (1st subscale) and the intellectual engagement processes (2nd subscale).

To establish the scale for measuring moderating functions three rating questions are used:

- **Opening comment** – Almost exclusively the teacher/tutor's responsibility; it means to open a discussion announcing the topic(s), contents, goals, etc.
- **Setting discussion norms and agenda –** Generally these aspects are the teacher's responsibility, and have to do with organization, calendarization, delegation and assessment.
- **Referring materials –** Indication of referring materials, bibliography, etc. that support learning.

The answers to these three rating questions allow to classify the individual – student or teacher/tutor – according to a scale of use of moderating functions that comprises five levels: Basic, Effective, Active, Strong and Expert, for the student; and Minimal, Basic, Effective, Strong and Expert, for the teacher/tutor.

Unidimensional Scale – Measuring the Quality of Engaged Collaborative Discourse

The three subscales previously presented set up the basis to build a Unidimensional Scale that allows the measuring of the Quality of Engaged Collaborative Discourse, using a set of interpretative anchors for the interactions that occurred. But according to Xin [13], to use all the possible combinations they offer, ordering them from low to high, would result in an overly complex and confusing overall scale of no practical use. On the

one hand, the communication, intellectual engagement, and use of moderating functions are closely related, so inevitably there are some overlaps among the levels of the three subscales. On the other hand, the combinations work at a theoretical level but might not translate into real life settings in one-to-one terms. Therefore, the author proposes a scale to measure the Quality of ECD comprising six levels: Peripheral, Minimal, Basic, Normal, Competent and Excellent. This scale is considered to be unidimensional and hypothetical, but provides an interpretative framework that allows the drawing of a progression map of the individual's performance in a given online course adopting a collaborative learning approach.

IV. Results: The Application of the Subscales to a Discussion Forum

As we have already mentioned, we used Xin's Scale of Engaged Collaborative Discourse [13] to analyze the discussion forum of a one-week module containing 52 messages.

1st Subscale

Table 1 – Measuring the Communication Processes

Trainee's Code	Login activity		Contribution level				Interactivity level			
	Yes	No	N	O	R	A	N	O	R	A
FSM	----	----	0	0	0	10	0	0	0	7
EC	----	----	0	0	0	8	0	0	6	0
LL	----	----	0	0	5	0	0	3	0	0
MM	----	----	0	2	0	0	0	1	0	0
VR	----	----	0	2	0	0	0	0	0	0
FS	----	----	0	2	0	0	0	0	0	0
BS	----	----	0	0	0	7	0	0	5	0
LP	----	----	0	0	6	0	0	0	6	0
LS	----	----	0	0	5	0	0	0	4	0
MG	----	----	0	1	0	0	0	1	0	0
FMJ	----	----	0	1	0	0	0	1	0	0
LJ	----	----	0	2	0	0	0	1	0	0
LN	----	----	0	3	0	0	0	1	0	0
TOTAL	----	----	0	13	16	25	0	8	21	7

Legend: N (None); **O** (Occasionally); **R** (Regularly); **A** (Actively)

Our intention to apply this scale met its first drawback when we realized that the functions that would allow us to quantify the login and download

activity of the individuals in the discussion forum were not implemented in the system used to deliver the course. This obviously prevented us from assessing the quantity of login activity.

We have considered as *Quantity of Contribution* the number of messages posted by a given individual in that discussion forum, dividing the total amount of 52 interventions by the participants. This process revealed systematically the participants that contributed the most to the discussion and those that did not even take part in it.

We have taken as *Quantity of Interaction* the number of messages related to one or more previous contributions, in a merely quantitative perspective, which equals to saying that we counted the branches in the discussion tree ignoring the contents of the contributions.

The attempt to apply this first subscale met yet another important drawback. Although some of the contributions were inserted as replies to other messages, they didn't reveal any indicator of relation with one or more previous messages after content analysis had been performed. This makes it obvious that often a quantitative analysis without articulation with a qualitative approach distorts the validity of the data obtained.

2nd Subscale

This 2nd subscale – Measuring Intellectual Engagement – has a qualitative nature, which implied an analysis of the meaning of the transcriptions.

Table 2 – Measuring Intellectual Engagement

Trainee's code	Coherent w/ teacher's agenda		New ideas and new points of view			Interactivity level		
	Yes	No	N/I	Y/L	YS	N/I	Y/1	Y/M
FSM	5	5	6	4	0	4	6	0
EC	4	4	3	5	0	4	3	1
LL	3	1	4	0	0	1	1	2
MM	1	1	2	0	0	1	1	0
VR	1	0	1	0	0	1	0	0
FS	2	0	2	0	0	2	0	0
BS	4	2	5	1	0	3	2	1
LP	0	6	6	0	0	3	3	0
LS	2	2	3	1	0	1	3	0
MG	0	1	1	0	0	0	0	1
FMJ	1	0	1	0	0	0	0	1
LJ	1	0	1	0	0	1	0	0
LN	1	1	2	0	0	1	1	0
TOTAL	25	23	37	11	0	22	20	6

Legend: N/I(None or Insignificant); **Y/L** (Yes but with little supporting arguments); **Y/S** (Yes with strong supporting arguments); **Y/1** (Yes with relating to one previous contribution); **Y/M** (Yes with relating to multiple previous contributions)

The fact is that the analysis of transcriptions of written asynchronous discussions is a complex process. Many researchers describe the processes of revision of these transcriptions as "content analysis", following a long tradition developed in the field of Communication Studies.

The different research domains of the various authors lead to different perspectives on content analysis, which has caused a lot of controversy since the 1940s. A linguist approaches a text with a wealth of theoretical resources which are quite different from those of a sociologist, psychologist or political science researcher. Each of them has their own theoretical references and different goals for the socially oriented studies.

According to Vala [11], it's the articulation between form and content that seems to offer a greater resistance, and content analysis has never been able to solve this problem. In his view, content analysis has mostly analyzed content items, sometimes making an effort to analyze stylistic procedures, but has never come up with a common framework for these two levels of analysis.

Berelson, (cit. Rourke [10:3]), on the other hand, defines content analysis as "a research technique for the objective, systematic, quantitative description of the manifest content of communication".

Following this definition, we sought to elaborate a categorization system that fitted the framework of the content analysis methodology proposed by the author according to its characteristics: objectivity, reliability, replicability, quantitative, description and inferential were the characteristics we favored.

Taking into account Vala's perspective [11:103] that "as a research technique, content analysis demands the highest specification possible of the procedures used", the categorization system was based on an interjudge agreement among the researchers involved in the study.

In fact, the discussion forum analyzed was focused on a lively discussion among the trainees about a text provided by the trainer. This aroused the need to elaborate also a content analysis grid for the text being discussed, since it was the only way in which we could classify the participants contributions according to Xin's proposition concerning new ideas and/or new points of view.

The application of Xin's 2nd subscale raised another methodological issue that we overcame, although aware that the process was becoming a highly subjective one – the reading of the scale grid and its possible articulation with the intellectual engagement levels defined by the author.

3rd Subscale

Finally, when we tried to apply the 3rd scale, Measuring Moderating Functions in ECD, and since the discussion was not subdivided to allow for small group work, the results obtained indicate that the trainer was the only participant to assume moderation functions. In a more cooperative-oriented task, an intragroup analysis would certainly result interesting.

Table 3 – Measuring Moderation Functions

Trainee's code	**Opening comment?**		**Setting norms?**		**Setting agenda?**		**Referring materials**	
	Yes	No	Yes	No	Yes	No	WO/E	W/E
FSM	2	0	6	0	4	0	2	3
EC	0	0	2	0	0	0	0	1
LL	0	0	1	0	0	0	0	0
MM	0	0	0	0	0	0	0	0
VR	0	0	0	0	0	0	0	0
FS	0	0	0	0	0	0	0	0
BS	0	0	0	0	0	0	0	0
LP	0	0	1	0	0	0	0	0
LS	0	0	0	0	0	0	0	0
MG	0	0	1	0	1	0	0	0
FMJ	0	0	0	0	0	0	0	0
LJ	0	0	0	0	0	0	0	0
LN	0	0	0	0	0	0	0	0
TOTAL	2	0	11	0	1	0	0	0

Legend: WO/E (Without Explanation); **W/E** (With Explanation);

When we combined the three subscales in search for a result that measured the quality of engaged collaborative discourse, we actually obtained an individual indicator of each participant's performance, although conditioned by the aforementioned methodological constraints. Notwithstanding, the application of Xin's three subscales didn't allow us to analyze the engaged collaborative discourse among the participants as a collaborative learning group.

Final Considerations

The study presented here is still an ongoing one. The subject of computer mediated collaborative learning and the level of engaged collaborative discourse on the part of the participants are highly relevant for our research.

Our use of Xin's subscales was not intended to describe rigorously the quality of engagement and produce results accordingly; our aim was to test this methodology, mapping its fragilities when applied to the analysis of asynchronous interactions in a discussion forum.

The classification proposed by Xin for the results obtained in the different subscales and at various levels of performance strikes us as being somewhat subjective. Since the results are organized in a qualitatively ascending order, without clearly specifying the criteria for the positioning in one or the other subsequent level, it is hard for another researcher to apply these procedures in a different context.

Even resorting to intersubjective processes to perform content analysis, it is consensual among many researchers the difficult, frustrating and subjective nature of this technique when trying to extract meaning from transcripts of asynchronous, computer mediated conferencing in formal educational settings.

It is nonetheless apparent that researchers have progressively gained experience with its use, which has reflected itself upon the evolution that the concept has undergone throughout time.

Riffe, Fico & Lacy, (cit. Rourke [10]), for example, describing the concept of content analysis four decades after Berelson, state that

> "quantitative content analysis is the systematic and replicable examination of symbols of communication, which have been assigned numeric values according to valid measurement rules, and the analysis of relationships involving those values using statistical methods, in order to describe the communication, draw inferences about its meaning, or infer from the communication to its context, both of production and consumption" (p. 20).

There is still a long way to go in the analysis of computer mediated communication if satisfactory levels of reliability, validity and replicability are to be attained in the analysis performed by different researchers. Only valid measurements can support assertions on the nature of asynchronous dialogue and its potential to generate insightful and sound arguments and discussions, in other words, on their true pedagogic potential.

Valid measurements certainly presuppose relevant categories of analysis. In that sense, we feel that Xin's model, although it may require necessary contextualizations and improvements, might serve as an excellent point of departure to establish solid methodological foundations for the analysis of online educational interactions.

Acknowledgements

The authors want to thank the collaboration provided by all the teachers and tutors involved in this study, as well as the coordinators of the respective courses.

References

1. Dillenbourg P, Baker M, Blaye A, O'Malley C (1995) The evolution of research on collaborative learning. In: Reiman P, Spada H (eds) Learning in humans and machines: towards an interdisciplinary learning science. Elsevier, Oxford, pp 189- 211
2. Fahy PJ (2001) Addressing some common problems in transcript analysis. International Review of Research in Open and Distance Learning, 1:1-5
3. Henri F (1992) Computer conferencing and content analysis. In Kaye A Collaborative Learning Through Computer Conferencing. Heidelberg The Najaden Papers NATO ASI Series Springer-Verlag, pp 117-136
4. Levy P (1993) The technologies of intelligence (in Portuguese). Editora 34, Rio de Janeiro
5. Lévy P (1999) Cyberculture (in Portuguese). Instituto Piaget, Lisbon
6. Niza S (1997) Cooperated trainning (in Portuguese). Educa, Lisbon
7. Panitz T Collaborative versus cooperative learning - a comparison of the two concepts which will help us understand the underlying nature of interactive learning. Available in: http://home.capecod.net/~tpanitz/tedsarticles/coopdefinition.htm 28/06/2006
8. Perrot-Clermont AN (1978) Intelligence development and social interaction (in Portuguese). Instituto Piaget, Lisbon
9. Piaget J (1973) Explanation in sociology. Sociological studies (in Portuguese). Companhia Editora Forense, Rio de Janeiro, pp 17-113
10. Rourke L, Anderson T, Garrison R, Archer W (2001) Methodological issues in analyses of asynchronous, text-based computer conferencing Transcripts. Available at http://communitiesofinquiry.com/documents/MethPaperFinal.pdf 28/06/2006
11. Vala J (1986) Content analysis (in Portuguese). In: Pinto M, Silva JS, Metodologia das Ciências Sociais. Edições Afrontamento, Porto
12. Vygotsky LS (1978) Mind in society: The development of higher psychological processes. MA: Harvard University Press, Cambridge
13. Xin C (2002) Validity centered design for the domain of engaged collaborative discourse in computer conferencing. Ph.D. thesis, Brigham Young University.